Article Watch: December 2019

Clive A. Slaughter

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, MCG-UGA Medical Partnership, 1425 Prince Ave., Athens, GA 30606, USA. Tel: (706) 713-2216; Fax: (706) 713-2221; E-mail: cslaught@uga.edu, or to any member of the editorial board. Article summaries reflect the reviewer's opinions and not necessarily those of the Association.

MACROMOLECULAR SYNTHESIS AND SYNTHETIC BIOLOGY

Wang K, de la Torre D, Robertson W E, Chin J W. Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly. *Science* 365;2019:922-926.

Wang *et al.* demonstrate a series of procedures for precise, large-scale, modular assembly of synthetic genomes on a megabase scale. First, they describe a general and efficient procedure for single-step, programmed splitting of the circular genome of *Escherichia coli* into two chromosomes. The procedure employs CRISPR-Cas9 for excision of a long piece of DNA from the initial *E. coli* chromosome and employs λ red recombinase for recircularizing the cut chromosome minus the portion excised. The excised piece is incorporated into a previously introduced bacterial artificial chromosome (BAC) by cutting with CRISPR-Cas9 and recircularizing with λ red recombinase. The process is accomplished with defined breakpoints and without sequence scars. Second, the authors perform programmed fusion of such synthetic chromosome pairs to yield defined genomic inversions and translocations. Third, they demonstrate chromosome transplantation from one strain to another, followed by chromosome fusion into a single genome for the construction of synthetic genomes from different progenitors. This methodology is anticipated to have a significant impact on the ability to rewrite bacterial genomes.

MASS SPECTROMETRY

Giles K, Ujma J, Wildgoose J, Pringle S, Richardson K, Langridge D, Green M. A cyclic ion mobility-mass spectrometry system. *Analytical Chemistry* 91; 2019:8564-8573.

A new commercial instrument for ion mobility-mass spectrometry (IM-MS) is described. The instrument incorporates a cyclic, closed-loop ion mobility sector. Ions carried around the 98-cm path-length loop by means of a traveling wave are allowed to traverse the loop multiple times under user control. The resolving power increases with the square root of the number of passes. With a single pass, resolution is ~80, and with 100 passes, resolution reaches ~750. The instrument, manufactured by Waters Corporation (Wilmslow, United Kingdom), has a quadrupole/IM/time-of-flight geometry. The IM loop is positioned orthogonally to the main optical axis of the instrument. A planar array of electrodes controls ion entry to and exit from the IM loop under operator programming. The instrument's capability is illustrated by separation of 3 isomeric pentasaccharides and the 6+ ion of the small protein ubiquitin. Collisional cross section of ions can be measured against an appropriate calibrant. The instrument incorporates the capability to program sequences of ion selection, activation, or fragmentation in the loop in a manner analogous to MS^n experiments.

FUNCTIONAL GENOMICS AND PROTEOMICS

Strecker J, Ladha A, Gardner Z, Schmid-Burgk J L, Makarova K S, Koonin E V, Zhang F. RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365;2019:48-53.

Klompe S E, Vo P L H, Halpin-Healy T S, Sternberg S H. Transposon-encoded CRISPR–Cas systems direct RNA-guided DNA integration. *Nature* 571; 2019:219-225.

The insertion of new DNA by conventionally implemented CRISPR-Cas utilizes the recipient cell's DNA repair mechanisms. Homology-directed repair and nonhomologous end joining are initiated by a double-stranded break made by the Cas nuclease. Although the site of the DNA breakage, programmed by guide RNA, is often highly specific, the introduction of insertions or

doi: 10.7171/jbt.19-3004-003

deletions (indels) during end joining may be problematic. Two groups now describe convenient transposon-based methodology for DNA insertion that does not invoke endogenous repair mechanisms. Their work derives from the earlier discovery of certain Tn7-like transposons that are associated with elements of the CRISPR machinery. These transposons mediate DNA binding without cleavage and insert transposon DNA into bacterial genomes at sites directed by CRISPR RNA (crRNA). Strecker *et al.* use the cleavage-incompetent Cas12k together with 3 transposase subunits (TnsB, TnsC, and TniQ) for insertion of a donor transposon into the *E. coli* genome. Insertion occurs 60–66 bp downstream of the site at which the small guide RNA binds and always occurs with the same orientation. This system, named CRISPR-associated transposase (CAST), is not entirely scarless. A 5-bp sequence before the insertion site is duplicated at the other end of the insert. This is concordant with the staggered DNA breaks generated by Tn7. Insertion rates vary with transposon size. Strecker *et al.* record ~60% insertion frequency with a 500-bp insert and ~25% with a 10-kbp insert. Insertion frequencies reached 80% at the most-favorable target sites, even without positive selection. However, off-target insertion approached 50%. Klompe *et al.* use a similar system. They employ an RNA-guided CRISPR complex called Cascade, which encodes Cas6, Cas7, and Cas8. Transposase subunits TnsA, TnsB, TnsC, and TniQ are supplied. Cascade binds to the transposase subunit TniQ, which mediates transposon insertion into the *E. coli* genome 48–50 bp downstream of the crRNA binding site. Insertion in both orientations occurs. Again, a downstream 5-bp duplication is present. Integration is at its optimal frequency with a transposon of 775 bp and decreases with shorter and longer pieces. Frequencies as high as 62% are recorded without positive selection. Target site specificity is very high (99%). These transposase systems must now be tested in eukaryotic cells. Meanwhile, they are raising hope for realizing a system for DNA insertion independent of endogenous DNA repair.

Lee D-S, Luo C, Zhou J, Chandran S, Rivkin A, Bartlett A, Nery J R, Fitzpatrick C, O'Connor C, Dixon J R, Ecker J R. Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nature Methods* 16;2019:999-1006.

Chromosome architecture and methylation pattern are both important features of chromatin organization that contribute to gene regulation. Chromosome architecture may be analyzed by 3C or HiC, and methylation patterns may be determined by bisulfite sequencing or related techniques. Lee *et al.* develop methodology to perform joint chromatin conformation and methylation profiling to assess the role each kind of analysis may play in discriminating different cell types at the single-cell level. They perform restriction enzyme digestion and DNA ligation on fixed nuclei for 3C analysis and then use fluorescence-activated sorting to dispense the treated nuclei individually into different wells in 384-well PCR plates. The nuclei are then subjected to proteinase digestion and bisulfite conversion for methylome sequencing by previously published methods (Luo *et al. Science* 357;2017:600-604). The authors validate the methodology by study of mouse embryonic stem cells (mESCs) and apply it in a study of >4200 cells of the human brain prefrontal cortex. They distinguish 14 different cell types. Their results indicate that 3C alone can distinguish cells of very different type, such as mESCs and murine mammary gland (NMuMG) cells, or neurons and nonneuronal cells in the brain. However, 3C does not reliably discriminate neuronal subtypes other than excitatory and inhibitory cells. The data further highlight genome-wide association between chromatin conformation and methylation in the control of gene expression.

Wang Y, Wang A, Liu Z, Thurman A L, Powers L S, Zou M, Zhao Y, Hefel A, Li Y, Zabner J, Au K F. Single-molecule long-read sequencing reveals the chromatin basis of gene expression. *Genome Research* 29;2019:1329-1342.

Wang *et al.* determine chromatin accessibility and nucleosome occupancy at the single-cell level using a combination of methyltransferase treatment and single-molecule long-read sequencing. Chromatin is treated with an exogenous methyltransferase to methylate cytosine to 5-methylcytosone (5mC) at sites not protected by nucleosome binding. Single-molecule sequencing is performed with a GridION sequencer from Oxford Nanopore Technologies (Oxford, United Kingdom). In nanopore sequencing, 5mC may be detected directly by its electrolytic current signal without bisulfite conversion. PCR amplification is unnecessary. The resulting ultralong reads span multiple genomic features and therefore assist in phasing. The authors apply these methods to haploid *Saccharomyces cerevisiae* cells. The single DNA molecules in this system represent single cells, so heterogeneity between single cells is directly amenable to study. Combining the results with single-cell RNA-seq permits quantitative investigation of the relationship between chromatin accessibility and gene transcription. The results illustrate the differential placement of nucleosomes surrounding transcriptional start sites in active

and inactive genes, the accessibility of promoters, and coordinate changes in accessibility of adjacent genes under regulatory control. This study provides methodology for investigation of chromatin on a single-cell basis and represents application of nanopore sequencing to epigenetics of single DNA molecules.

PROTEOMICS

Brenes A, Hukelmann J, Bensaddek D, Lamond A I. Multibatch TMT reveals false positives, batch effects and missing values. *Molecular & Cellular Proteomics* 18;2019:1967-1980.

Isobaric labeling strategies have enabled multiplexed analysis of samples for which both identification and quantification is needed. Both tasks can be performed by data-dependent liquid chromatography-tandem mass spectrometry ("shotgun" proteomics) of the labeled peptides. Despite the inherently stochastic sampling of peptides for fragmentation in such analyses, multiplexing has the additional advantage of minimizing the "missing values" problem. Nevertheless, biologic interpretation of quantitative measurements requires replication. Using the 10-plex tandem mass tag (TMT) system to analyze induced pluripotent stem cell (iPSC) samples, Brenes *et al.* set out to determine the effect of replication on the number of missing values encountered when comparing independent multiplex TMT analyses. Surprisingly, when 2 analyses are compared, the within-analysis number of missing values for peptides, $<2\%$, rises to 24%, and for proteins it rises from $<0.5\%$ to $>6\%$. This problem substantially impairs the confidence with which results can be interpreted. The authors also assess reproducibility of protein quantification between multiplexed analyses. The coefficient of variation of protein copy number between technical replicates of multiplexed analyses was 6.4-fold higher than the coefficient of variation within a single multiplexed analysis. The authors therefore advocate inclusion of an internal standard sample of suitable peptide diversity for purposes of normalization when comparing different 10-plex TMT batches. They also assess the rate of false peptide identifications ("false positives"). The donors of iSPC cells are both male and female. The cells from females should lack peptides derived from proteins encoded by the Y chromosome. However, in a 10-plex batch containing only 2 male channels, the remaining 8 female channels still quantified 97.5% of all Y chromosome–specific peptides, presumably because of peptide coisolation interference. Reporter ion interference was also noted for these peptides. This may produce distortion of quantitative measurements. The results of this study highlight limitations in multiplexed analyses that should be considered when designing quantitative proteomic studies.

CELL BIOLOGY AND TISSUE ENGINEERING

Ombrato L, Nolan E, Kurelac I, Mavousian A, Bridgeman V L, Heinze I, Chakravarty P, Horswell S, Gonzalez-Gualda E, Matacchione G, Weston A, Kirkpatrick J, Husain E, Speirs V, Collinson L, Ori A, Lee J-H, Malanchi I. Metastatic-niche labeling reveals parenchymal cells with stem features. *Nature* 572;2019:603-608.

Metastatic cancer cells may die or lay dormant in destination tissues for decades before forming tumors. The characteristics of the "metastatic niche" that govern these behaviors are largely unknown. Ombrato *et al.* provide new methodology with which to address this issue. In a mouse model of metastatic breast cancer, they perform intravenous injection of breast cancer cells that express a fluorescent mCherry protein engineered to incorporate a membrane-permeable peptide. In the destination tissue, the mCherry protein diffuses out of the cancer cells. It reenters the secreting cells and also enters neighboring parenchymal cells to a depth of about 5 surrounding cell layers. Fluorescence-activated cell sorting (FACS) then enables identification and characterization of cells making up the metastatic niche. Cellular changes induced in the proximity of cancer cells may therefore be detected. In studies of metastasis of breast cancer cells to the lung, the authors identify alveolar type II (AT2) cells and CD45+ immune cells as the main cancer-associated parenchymal cells. The AT2 cells may adopt an activated state with acquisition of stem cell–like features including expression of lung progenitor markers and potential for multilineage differentiation. The mCherry system is expected to contribute significantly to our understanding of metastasis.

Zheng Y, Xue X, Shao Y, Wang S, Esfahani S N, Li Z, Muncie J M, Lakins J N, Weaver V M, Gumucio D L, Fu J. Controlled modeling of human epiblast and amnion development using stem cells. *Nature* 573; 2019:421-425.

A microfluidic device is described for *in vitro* culture of human pluripotent stem cells (hPSCs) that supports *in vitro* recapitulation of developmental landmarks in post-implantation human embryos. The device contains 3

parallel channels: a central, matrix-filled channel with posts 80 μm apart that demarcate matrix pouches within which hPSCs can develop, a cell loading channel from which hPSCs migrate into the matrix, and a channel carrying culture medium for controlled introduction of growth factors, morphogens, and/or antagonists to cells in the matrix. The model exhibits spatial coordination of cell differentiation events. Within 36 h, cells undergo development of an epiblast-like cyst, within which a lumen analogous to the proamniotic cavity forms spontaneously. The cyst is asymmetric, with amniotic ectoderm-like cells representing a prospective dorsal pole. With the introduction of particular morphogens, this cyst reproducibly forms an anteriorized or a posteriorized embryonic-like sac analogous to the anterior and posterior ends of an embryo. Primitive streak-like cells and primordial germ-like cells are detected. This model system is hoped to contribute to knowledge of developmental events in the postimplantation human embryo.

IMAGING

Saka S K, Wang Y, Kishi J Y, Zhu A, Zeng Y, Xie W, Kirli K, Yapp C, Cicconet M, Beliveau B J, Lapan S W, Yin S, Lin M, Boyden E S, Kaeser P S, Pihan G, Church G M, Yin P. Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nature Biotechnology* 37;2019:1080-1090.

Saka *et al.* demonstrate methodology for multiplexed immunostaining of diverse proteins in tissue samples with sensitivity levels individually programmable to accommodate a large range of abundance of different target proteins and with high throughput for rapidly staining multiple proteins in the sample. The sample is first stained simultaneously with primary DNA-barcoded antibodies with specificity for the multiple protein targets to be localized. The barcodes are then hybridized, again simultaneously, to orthogonal, single-stranded DNA strands that additionally bear concatemers with predetermined numbers of repeats. A fluorophore-labeled imager strand is then hybridized to the repeated concatemer binding sites, with fluorescence intensity determined by the number of repeats. Multiple targets are imaged sequentially by hybridization and dehybridization of different orthogonal imaging strands in multiple, rapid exchange cycles. The concatemers are synthesized in primer exchange reactions that employ a strand-displacing polymerase. The authors demonstrate localization of a variety of proteins and application to diverse types of sample, including cultured cells, cryosections, formalin-fixed and paraffin-embedded sections, and whole-mount tissues. They achieve signal amplification from 5- to 180-fold, matching or exceeding sensitivities attained by conventional techniques. They demonstrate serial imaging of 10 protein targets at high spatial resolution. The research group has previously shown that similar methodology is suitable for FISH to localize DNA and RNA (Kashi *et al. Nature Methods* 16; 2019:533-544). The authors anticipate application in development of tissue atlases, profiling of pathologic tissues, and biomarker discovery and screening, among other applications.

Barré F P Y, Paine M R L, Flinders B, Trevitt A J, Kelly P D, Ait-Belkacem R, Garcia J P, Creemers L B, Stauber J, Vreeken R J, Cillero-Pastor B, Ellis S R, Heeren R M A. Enhanced sensitivity using MALDI imaging coupled with laser postionization (MALDI-2) for pharmaceutical research. *Analytical Chemistry* 91;2019:10840-10848.

Niehaus M, Soltwisch J, Belov M E, Dreisewerd K. Transmission-mode MALDI-2 mass spectrometry imaging of cells and tissues at subcellular resolution. *Nature Methods* 16;2019:925-931.

Two groups exploit the enhanced ion yields in mass spectral imaging when plasma ionization (so-called postionization) is coupled to matrix-assisted laser desorption/ionization (MALDI) at atmospheric pressure. In this technique, particles are initially desorbed by pulsed illumination from a UV laser from a sample embedded in a UV-absorbing matrix, as in conventional MALDI. However, photons from a UV laser are additionally directed at the evolving MALDI plume of ions and neutral species (in a process known as MALDI-2). This causes ion/molecule interactions that increase ion yield, and therefore sensitivity, up to 2 orders of magnitude. Barré *et al.* use this technique for analysis and imaging the distribution of pharmaceutical compounds in human and animal tissues. Detection of drug molecules is normally difficult because drug concentrations in tissues may be very low, drug molecules may possess structures unfavorable for detection by MALDI, or the signals may be obscured by matrix interference. Detection is demonstrated in this study without the need for *in situ* derivatization to boost ion yields. Niehaus *et al.* perform postionization to enable a decrease in MALDI pixel size. The smaller the area of tissue from which ions are sampled, the lower the signal strength. This study achieves brain tissue imaging with a pixel size of just 600 nm. This capability will enhance image resolution, which is hoped to enable subcellular mass spectral imaging by MALDI.

DRUG DEVELOPMENT

Johnson E O, LaVerriere E, Office E, Stanley M, Meyer E, Kawate T, Gomez J E, Audette R E, Bandyopadhyay N, Betancourt N, Delano K, Da Silva I, Davis J, Gallo C, Gardner M, Golas A J, Guinn K M, Kennedy S, Korn R, McConnell J A, Moss C E, Murphy K C, Nietupski R M, Papavinasasundaram K G, Pinkham J T, Pino P A, Proulx M K, Ruecker N, Song N, Thompson M, Trujillo C, Wakabayashi S, Wallach J B, Watson C, Ioerger T R, Lander E S, Hubbard B K, Serrano-Wu M H, Ehrt S, Fitzgerald M, Rubin E J, Sassetti C M, Schnappinger D, Hung D T. Large-scale chemical–genetics yields new *M. tuberculosis* inhibitor classes. *Nature* 571; 2019:72-78.

Screens for antibiotic discovery that are based on biochemical assays of individual target proteins may identify compounds that lack the ability to kill pathogen cells. Conversely, screens that are based on whole-cell assays may identify compounds with a mechanism of action that is difficult to determine. In the context of a program to discover new antibiotics against *Mycobacterium tuberculosis*, the authors develop a strategy that combines strengths of both these approaches. They screen libraries of compounds against a population of mycobacterial cells comprising a pool of "hypomorphic" cell clones. Hypomorphic cells are cells with depleted levels of expression of 1 or other essential genes, numbering 474 in this study. The depletion of target proteins is achieved here by accelerating the rate of their degradation *via* addition of a C-terminal tag that targets the protein for intracellular proteolysis. The degree of depletion is adjusted to avoid impairing the cell's growth rate, but the cell is rendered susceptible to any further inhibition by any of the compounds in the library with which the target protein interacts. Each cell clone is further tagged with a 20-nt DNA barcode to track depletion of cells in the pool by sequence-based barcode counting. Such an extensive panel of hypomorphic cells enables a much larger number of potential target genes to be tested than in studies restricted to the proteins targeted by existing drugs. The hypomorphs' sensitivity to growth inhibition results in 10-fold more hits than by screening against the wild-type *Mycobacterium*. The number of affected clones is used as a proxy for the specificity of any particular compound in the tested panel, and their identity provides information about mechanism of action. The authors identify the mechanism of action of 45 new antimicrobial compounds. These include an inhibitor of a new target protein, the efflux pump EfpA, as well as new molecular scaffolds interacting with established targets.

Zimmermann M, Zimmermann-Kogadeeva M, Wegmann R, Goodman A L. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature* 570;2019:462-467.

It has long been known that commensal microbes in the gastrointestinal (GI) tract can metabolize certain orally administered drugs. Clinically significant effects include drug activation, inactivation, or toxification. In order to determine how pervasive chemical modification of drugs by the human GI microbiota might be, 76 human GI microbial species/strains representing the major phyla of the human gut microbiome are tested for their ability to metabolize 271 drugs. The chosen drugs sample diverse clinical indications (although not antibiotics), physicochemical characteristics, and predicted concentrations in the GI tract. Drug concentrations are measured before and after a 12-h incubation with bacteria in aerobic culture, and metabolic products are identified using liquid chromatography-mass spectrometry (LC-MS) for untargeted metabolomic analysis. The authors find that 176 of the drugs undergo metabolism by at least 1 bacterium, causing significant reduction in the level of the drug. Every bacterium metabolizes some of the drugs, with numbers of drugs ranging from 11 to 95 per strain. Chemical changes included oxidation, reduction, acetylation, deacetylation, hydrogenation, and propionylation. Some drugs produced more than one metabolic product. The authors identify some of the enzymes responsible in a study of one exemplary bacterial species, *Bacteroides thetaiotaomicron*, which metabolizes 46 of the tested drugs. A gain-of-function screen is conducted in which 2–8-kb fragments of genomic DNA from this bacterium are cloned into *E. coli* using an expression vector. The transformants are tested for their ability to metabolize target drugs. Thirty microbial enzymes are identified that together convert 20 drugs to 59 metabolites. The study establishes a surprisingly high incidence of drug metabolism by the GI microbiome and provides methods for investigating mechanisms of microbial metabolism of pharmaceuticals and for determining the contribution of GI microbiome to pharmacodynamic variation between persons.

Lin A, Giuliano C J, Palladino A, John K M, Abramowicz C, Yuan M L, Sausville E L, Lukow D A, Liu L, Chait A R, Galluzzo Z C, Tucker C, Sheltzer J M. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science Translational Medicine* 11;2019:eaaw8412.

There continues to be much interest in the development of anticancer therapeutic agents that target proteins "essential" for cancer cells, proteins on which survival or proliferation of cancer cells depend. Lin *et al.* investigate the target specificity of 10 small-molecule cancer drugs that target 6 proteins identified as essential based on RNA interference (RNAi). Each of these drugs exhibits potent cell killing *in vitro* and *in vivo* and is undergoing clinical or late-stage preclinical testing. Nonetheless, each drug lacks known mutations in its target protein that confer resistance to the target inhibitors. Such mutations would authenticate the inhibitors' specificity. The authors induce knock out of each gene by CRISPR-Cas9 mutagenesis. Remarkably, the knockout clones continue to proliferate at rates comparable to control cells. Experiments to eliminate the possibility that various adaptive mechanisms might have played a part in producing these results lead to the conclusion that RNAi promiscuity has instead contributed to the misidentification of these putative drug targets. But the efficacy of the drugs directed against them nonetheless remains undiminished by loss of their supposed targets. This indicates that cell killing is occurring by off-target effects. One drug, OTS964, which was designed as an inhibitor of PDZ-binding kinase (PDK), is shown actually to work by inhibiting cyclin-dependent kinase 11 (CDK11). Inhibitors of various CDKs, although not this one, have previously been identified, and some have received FDA approval for treatment of malignancies. Accurate identification of a drug's mechanism of action is clearly important in selecting settings in which it is likely to be effective. The authors therefore urge that the mechanism of action of cancer drugs be rigorously established in the preclinical phase of their testing.